

a) COMPLETE

	①	②	③	④	⑤
①	0	9	3	6	11
②	9	0	7	5	10
③	3	7	0	9	2
④	6	5	9	0	8
⑤	11	10	2	8	0

• $d(3,5) = 2 \rightarrow$ AGRUPO (3,5) NA ALTURA 2

• BUSCAR MENOR DISTANCIA ENTRE $\{1, 2, (3,5), 4\}$

$\hookrightarrow d(1,2) = 9$

$\hookrightarrow d(1, (3,5)) = 11$ [MAIOR DISTANCIA]

$\hookrightarrow d(1,4) = 6$

$\hookrightarrow d(2, (3,5)) = 10$

$\hookrightarrow d(2,4) = 5$

$\hookrightarrow d((3,5), 4) = 9$

AGRUPO (2,4) NA ALTURA 5

• BUSCAR MENOR DISTANCIA ENTRE $\{1, (2,4), (3,5)\}$

$\hookrightarrow d(1, (2,4)) = 9$

$\hookrightarrow d(1, (3,5)) = 11$

$\hookrightarrow d((2,4), (3,5)) = \{7, 10, 9, 8\} = 10$

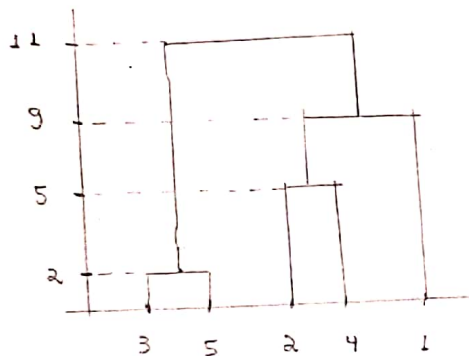
AGRUPO (1, 2, 4) NA ALTURA 9

• BUSCAR MENOR ENTRE $\{(1,2,4), (3,5)\}$

$\hookrightarrow d((1,2,4), (3,5)) = \{3, 11, 7, 10, 9, 8\} = 11$

AGRUPO (1,2,4,3,5) NA ALTURA 11

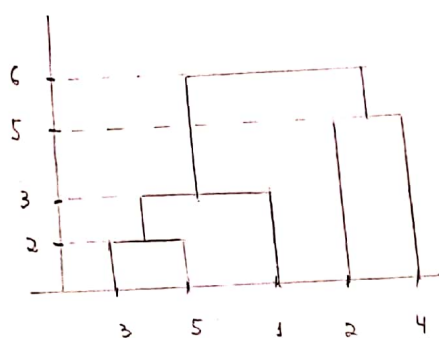
ESBOÇO DO DENDOGRAMA



b) SINGLE

	①	②	③	④	⑤
①	0	9	3	6	11
②	9	0	7	5	10
③	3	7	0	9	2
④	6	5	9	0	8
⑤	11	10	2	8	0

ESBOÇO DO DENDROGRAMA



$$d(3, 5) = 2$$

AGRUPO (3, 5) NA ALTURA 2

• BUSCAR MENOR DISTANCIA ENTRE $\{1, 2, (3, 5), 4\}$

$$d(1, 2) = 9$$

$$d(1, (3, 5)) = 3 \quad [\text{MENOR DISTANCIA}]$$

$$d(1, 4) = 6$$

$$d(2, (3, 5)) = 7$$

$$d(2, 4) = 5$$

$$d((3, 5), 4) = 8$$

AGRUPO (1, 3, 5) NA ALTURA 3

• BUSCAR MENOR DISTANCIA ENTRE $\{(1, 3, 5), 2, 4\}$

$$d((1, 3, 5), 2) = \{9, 7, 10\} = 7$$

$$d((1, 3, 5), 4) = \{6, 9, 8\} = 6$$

$$d(2, 4) = 5$$

AGRUPO (2, 4) NA ALTURA 5

• BUSCAR MENOR DISTANCIA ENTRE $\{(1, 3, 5), (2, 4)\}$

$$d((1, 3, 5), (2, 4)) = \{9, 6, 7, 9, 10, 8\} = 6$$

AGRUPO (1, 3, 5, 2, 4) NA ALTURA 6

c) Um aglomerado terá $\{3, 5\}$ e o outro terá $\{2, 4, 1\}$

d) Um aglomerado terá $\{3, 5, 1\}$ e o outro terá $\{2, 4\}$

(2) a)

ID	X_1 CHEST PAIN	X_2 SEX	X_3 SMOKES	X_4 EXERCISES	TARGET ↓ HEART ATTACK
1	1	1	0	1	1
2	1	1	1	0	1
3	0	0	1	0	1
4	0	1	0	1	0
5	1	0	1	1	1
6	0	1	1	1	0

P/ CHEST PAIN

$$G_0 = 3/6 \rightarrow \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} \rightarrow 1 - \left[\frac{1}{3} \right]^2 - \left[\frac{2}{3} \right]^2 = \frac{4}{9}$$

$$G_1 = 3/6 \rightarrow \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \rightarrow 1 - \left[\frac{3}{3} \right]^2 - \left[\frac{0}{3} \right]^2 = 0$$

$$G_{X_1} = \left(\frac{3}{6} \right) \cdot 0 + \left(\frac{3}{6} \right) \cdot \frac{4}{9} = \frac{2}{3} = \boxed{0,666}$$

P/ SEX

$$G_0 = 2/6 \rightarrow \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \rightarrow 1 - \left[\frac{2}{2} \right]^2 - \left[\frac{0}{2} \right]^2 = 0$$

$$G_1 = 4/6 \rightarrow \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} \rightarrow 1 - \left[\frac{2}{4} \right]^2 - \left[\frac{2}{4} \right]^2 = 0,5$$

$$G_{Y_2} = \left(\frac{2}{6} \right) \cdot 0 + \left(\frac{4}{6} \right) \cdot \frac{1}{2} = \frac{1}{3} = \boxed{0,33}$$

P/SMOKES

$$G_0 = 2/6 \rightarrow \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \rightarrow 1 - \left(\frac{1}{2}\right)^2 - \left(\frac{1}{2}\right)^2 = \frac{1}{2} = 0,5$$

$$G_1 = 4/6 \rightarrow \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \rightarrow 1 - \left(\frac{3}{4}\right)^2 - \left(\frac{1}{4}\right)^2 = \frac{3}{8} = 0,375$$

$$G_{X3} = \frac{2}{6} \cdot \frac{1}{2} + \frac{4}{6} \cdot \frac{3}{8} = \frac{5}{12} = \boxed{0,416}$$

P/ EXERCISE

$$G_0 = 2/6 \rightarrow \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \rightarrow 1 - \left(\frac{2}{3}\right)^2 - \left(\frac{0}{3}\right)^2 = 0$$

$$G_1 = 4/6 \rightarrow \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \rightarrow 1 - \left(\frac{2}{4}\right)^2 - \left(\frac{2}{4}\right)^2 = 0,5$$

$$G_{X4} = \frac{2}{6} \cdot 0 + \frac{4}{6} \cdot \frac{1}{2} = \frac{1}{3} = \boxed{0,33}$$

1ª ETAPA

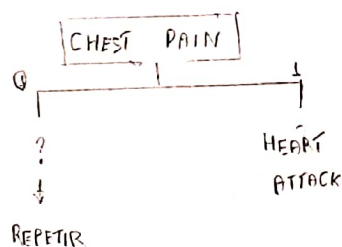
$$X_1 = 0,222$$

$$X_2 = 0,33$$

$$X_3 = 0,416$$

$$X_4 = 0,33$$

∴ CHEST PAIN
ESTARÁ NO 1º NO



NOVA TABELA

ID	CHEST PAIN	SEX	SMOKES	EXERCISE	HEART ATTACK
3	0	0	1	0	1
4	0	1	0	1	0
6	0	1	1	1	0

P/SEX

$$G_0 = 1 - \left[\frac{1}{1} \right]^2 - \left[\frac{0}{1} \right]^2 = 0$$

$$G_{X_3} = \frac{2}{3} \cdot 0 + \frac{1}{3} \cdot 0 = 0$$

$$G_1 = 1 - \left[\frac{0}{3} \right]^2 - \left[\frac{2}{2} \right]^2 = 0$$

P/SMOKES

$$G_0 = 1 - \left[\frac{0}{1} \right]^2 - \left[\frac{1}{1} \right]^2 = 0$$

$$G_{X_3} = \frac{2}{3} \cdot \frac{1}{2} + \frac{1}{3} \cdot 0 = \frac{1}{3} = 0,33$$

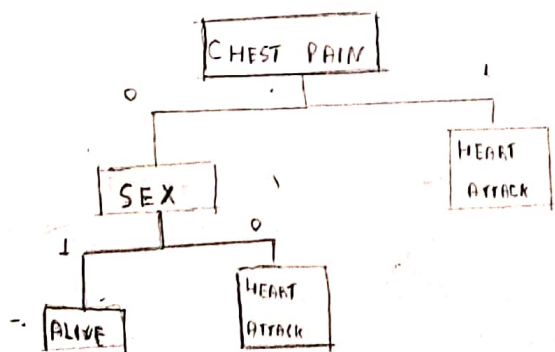
$$G_1 = 1 - \left[\frac{1}{2} \right]^2 - \left[\frac{1}{2} \right]^2 = \frac{1}{2}$$

P/EXERCISES

$$G_0 = 1 - \left[\frac{1}{1} \right]^2 - \left[\frac{0}{1} \right]^2 = 0$$

$$G_{X_4} = \frac{2}{3} \cdot 0 + \frac{1}{3} \cdot 0 = 0$$

$$G_1 = 1 - \left[\frac{0}{2} \right]^2 - \left[\frac{2}{2} \right]^2 = 0$$



- b) Com relação a espécie, o estudo infere que dentre elas a que tem a maior media de peso e desvio padrão é a Gentoo sendo muito diferente das outras duas, Chinstrap e Adelie, pois elas possuem quase o mesmo peso médio e desvio padrão, sendo a Chinstrap a maior se for levado em conta a media de peso médio e a Adelie a maior em desvio padrão. Um fato interessante é se pesquisarmos um grande grupo de penguins da espécie Gentoo com relação ao seu peso, veremos que há muitos penguins bem "magros" enquanto que de outra forma vemos penguins bem "gordos". Além disso, observando o comprimento das asas, vemos que com relação à média do comprimento a espécie Gentoo ainda fica em primeiro lugar, sendo seguido pela Chinstrap e depois pela Adelie, contudo levando em conta o desvio padrão de cada uma, ocorre uma mudança, na qual a Chinstrap acaba por ter o maior desvio padrão, sendo seguido pela Adelie e depois pelo Gentoo, um fato interessante que podemos tirar disso é que se observamos vários penguins da espécie Gentoo muitos terão as nadadeiras maiores e nenhum deles terá muito diferença com relação a elas, já se observarmos a Chinstrap, veremos que muitas tem nadadeiras pequenas, contudo a maioria dos penguins de sua espécie tem o comprimento das asa variando bastante então veremos penguins de asas bem pequenas e outras com asas bem grandes.
- c) Pelo estudo pode-se observar que os penguins machos tem maior peso médio e desvio padrão do que as fêmeas, além disso com relação ao comprimento da sua asa, os machos também possuem maior comprimento de asa porém menor desvio padrão do que as fêmeas, ou seja, as fêmeas quando levadas em conta em uma grande quantidade cada uma não se difere das outras por muita diferença, a maioria tem o mesmo peso e o mesmo comprimento de asa.
- f) Sim, há uma relação entre ambas as partes do penguim, como mostrado pelo plot. Ademais, em vista da análise feita pela função correlação pode-se aferir que a regressão linear entre as partes que mostra-se como positivas, ou seja, quando uma cresce, a outra também crescerá, além disso podemos afirmar que a correlação é forte pois é maior que 0.7, tendo o valor 0.8712 aproximadamente.
- h) O coeficiente angular da reta gerada pelo exercício "g" representa o aumento da massa do penguim com relação ao tamanho da asa podendo ser positiva, ou seja quando um cresce o outro também crescerá, ou podendo ser negativa que é o oposto. Ainda sim, um ponto importante de observar é que o aumento da asa em uma unidade de medida ("cm,mm ou m", depende de qual a pesquisa está usando), causa a adição igual ao seu coeficiente de variação no peso do penguim com relação ao seu peso anterior isso pode ser descrito pela formula:
 Seja, "a" o coeficiente de variação
 x a var que descreve o comp. da asa e y o que descreve o peso.
 $y_0 = ax_0 + b$
 X1 é a variável do aumento da asa que aumentou em 1 com relação ao seu anterior(X0)
 $y_1 = ax_1 + b \Leftrightarrow a(x_0+1) + b \Leftrightarrow y_1 = y_0 + a$

- i) Não, já que os dados limites dos dados são 172(mínimo) e 231(máximo) na reta estimada, portanto o valor 168 não pode ser estimado por meio desta reta, já que a função "lm" não levou em conta esse valor na hora de computar o coeficiente da reta e sua variável independente, sendo um número fora dos limites estabelecidos.
-
-